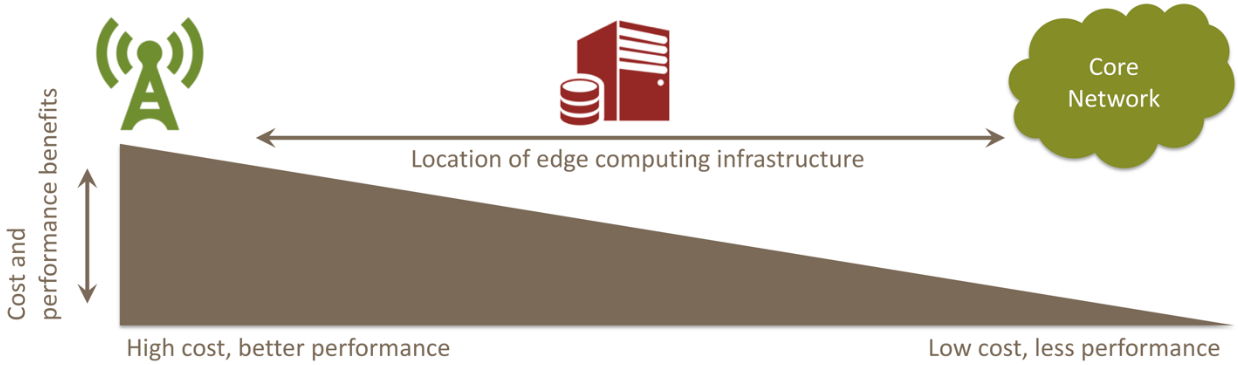
# The Wireless Edge

## Edge Computing in Wireless Networks




**Abstract:**

**Edge computing will add costs to the network but also offers revenue opportunities for operators and benefits for Private LTE networks. This report provides a summary of each initiative and the technical benefits, as well as the business case for edge computing in neutral host and operator scenarios. The plans of eight key operators for Edge Computing are listed. A five-year forecast is provided for mobile operator revenues enabled by Edge Computing, as well as a hardware/software forecast for Edge Computing equipment.**

**March 2018**

MOBILE EXPERTS

## Table of Contents

## List of Figures



## List of Tables

## Executive Summary

Edge Computing (EC) is an architecture that places compute and storage functions close to the edge of the network. The primary benefits of this approach are improved user experience through faster application response times and/or a reduced requirement for backhaul capacity.

At a business level, we believe that most consumer applications will be optimal not with Edge Computing at the radio site, but with computing at regional data centers that are located 10-100 miles away. The loss of ultra-low latency will impact revenue opportunities, but the savings in CAPEX for computing resources will be more important over the next five years.

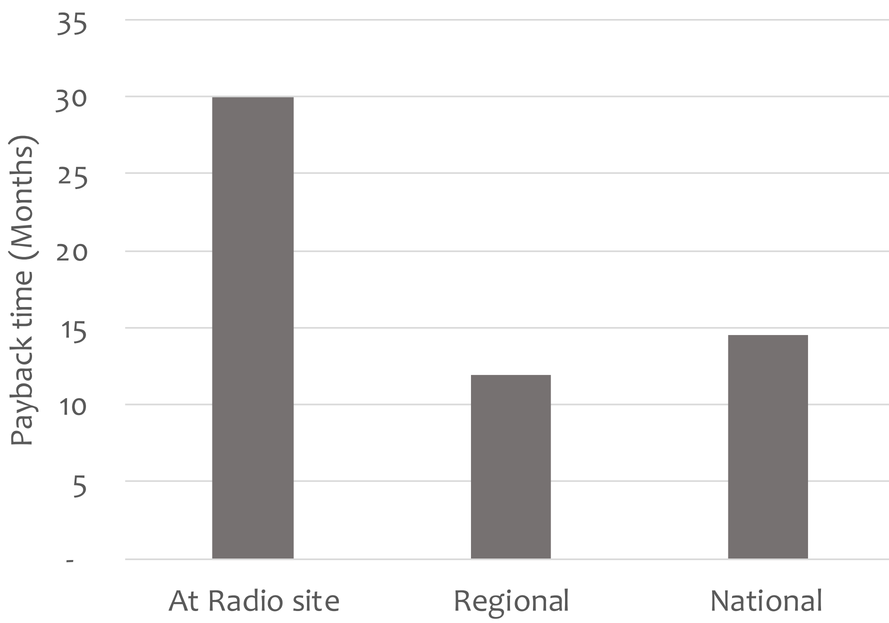



**Figure 1 ROI comparison of various Edge Computing locations.**

Enterprise applications may justify edge computing at the local radio site. The ultra-low latency requirements for robotics applications could justify the installation of computing platforms locally. In these specific cases, the payback period is extremely quick, and there's no need to install edge computing resources everywhere.

Our in-depth research on edge computing applications in wireless networks led to the following observations and conclusions:

1- Mobile network operators (MNOs) consider edge computing to be part of the 5G network architecture. Operators are not inclined to deploy edge computing at large scale with 4G networks.
2- Operators hesitant on placing edge computing between the core network and the access edge (base station) because of conflict with the 4G network architecture that centralizes billing and legal intercept functions in the core network.

3- The deployment of edge computing will follow closely the migration to virtualized networks which improves economics in favor of decentralizing the core network and pushing core network elements closer to the user – an architecture that's amenable to edge computing. AT&T and Verizon are favorable to this architecture.
4- Operators have expressed higher interest in services that allow them monetizing edge computing applications than in applications that allow them to save network operation costs alone, such as backhaul costs.
5- Operators are in process of defining the cost structure and market opportunity for edge computing. Almost all operators have told us they are not considering edge computing at the macrocell. This leaves three viable alternative locations: at the core network where the core is located close to the user, at the Cloud RAN baseband center (e.g. in central offices) and at the enterprise location where small cells are typically used.
6- Edge computing promises to provide mobile network operators a tool to compete with the OTTs. However, competition from OTTs promises to be fierce as they are in process of extending their Cloud services closer to the edge and are thus building a time-to-market market lead.
7- Edge computing is a key part of private networks. The CBRS shared spectrum regime and operation in unlicensed band (5 GHz) will positively impact the deployment of edge computing. This could lead to new business models where neutral hosts and MVNOs leverage edge computing to provide differentiated services complementing those of the mobile network operator.
8- MNOs need to take an IT-centric approach to align with the enterprise. Many MNOs will also have to create new sales channels into the enterprise. Enterprises on the other hand need to harmonize their IT infrastructure with their operational technology processes. These are among the market challenges impeding wider deployments of edge computing.
9- We forecast the cumulative revenue to service providers in the enterprise and consumer market segments from edge computing services to reach $7.2 billion through 2023. The United States will account for 59% of the revenues.
10- We forecast spending on edge computing to reach a cumulative $4.4 billion through 2023 to include hardware and software.

## Introduction

The mobile cloud is a successful architecture, but its viability is in question for future applications. Today, the 2.5 billion people who connect to the Internet through smartphones are served by the cloud, which is a relatively centralized interconnected network of computing power and storage aggregated in huge data centers around the globe. But this model has pitfalls in many different aspects. For instance, Internet of Things promises to connect billion of devices, not all of which need or could be connected to the Cloud. The time delay between devices and the Cloud, or latency, could be over the application requirements. The amount of data that devices generate is too high and costly to be transported to the network core or the Cloud. The requirements of new applications and use cases are forcing a rethinking of the mobile cloud architecture. The new architecture moves computing and storage resources closer to the devices and users. This is edge computing, a term that we use in this report irrespective of how the devices are connected to the network.

In this report, we will focus on the development of Edge Computing in wireless networks. Fundamentally, there are two main benefits for MNOs from edge computing. The approach that MNOs will use impacts the market size and business opportunity for vendors:

1. Opex reduction: Optimize network resource utilization through edge computing. This comes down to reducing amount of data and control signaling between the device and the core network. Edge computing for opex reduction represent a relatively smaller slice of the EC business opportunity.
2. Monetization of new services: edge computing helps MNOs provide new services with aggressive performance parameters compete with the OTTs and Cloud service providers by leveraging proximity to the end user. This represent the larger slice of the EC business opportunity, but it has higher level of uncertainty and risk due to the competitive environment.

Note that we specifically refrain from using the term “Mobile Edge Computing” as it evokes the connotation of the ETSI Multi-access Edge Compute project. ETSI MEC project began in addressing specific implementation topics that a few major ecosystem players did not concur with but has since broadened its scope.

## Methodology

To prepare this research we relied on first hand conversations and primary research with leading mobile MNOs, equipment vendors, system integrators, and a sample of large enterprises to gauge the status of edge computing in wireless networks and its future prospects. We also developed internal models factoring critical parameters on the readiness of the MNOs and enterprises to implement edge computing. Edge computing can only be considered in context of a number of developments in adjacent fields. For example, our research couples edge computing with network virtualization. Additionally, it couples enterprise deployments with the business models followed by MNOs, system integrators and vendors into the enterprise space. The enterprise space cannot be viewed in isolation from developments in heterogeneous networks including small cells and how MNOs approach that market.

## The Competitive Landscape: The Cloud Players

In analyzing the edge computing opportunity, we have to consider the competitive landscape where MNOs square off against the large Internet players. Over-the-top service providers (OTTs) leverage service agility and web-scale architecture to provide services leveraging economies of scale that MNOs operating at national and regional levels find it difficult to match. On the other hand, MNOs have the advantage of proximity to the end user. The proximity advantage is under threat if MNOs fail to capitalize on it, rendering them to “dumb pipes.” Edge computing is a vehicle to avoid that fate. Nevertheless, the financial industry values OTTs higher than MNOs assigning $1.9 trillion market capitalization[1] to the top 3 cloud companies, Amazon, Google and Microsoft, which is approximately 4x the market capitalization of the top 4 US telecom service providers: Verizon, AT&T, T-Mobile USA, and Sprint.

The cloud players recognize that the challenges associated with the distance between the centralized cloud (term used here in relative manner) and the device. They are racing towards the edge with new services that build upon and enhance existing enterprise services. Both AWS and Microsoft proceeded to launch their edge services, Greengrass and Azure Edge, respectively, in 2017. These services build upon the existing cloud business, augmenting it with additional edge services. Cloud players are agnostic of the access technology which gives focusing on the cloud implementation and value-added services provided to their customers, such as analytics and other services. They are also focusing on developed markets where industry is most likely to benefit from their services.

For additional information on this topic, see Appendix 1 for overview of AWS Greengrass.

In considering the MNOs opportunity in edge computing we factor the competition from the cloud players who have already jumped on this space to achieve a head start. The business model is another key aspect as cloud players are building on established presence in the enterprise market. This is a critical point as MNOs access to the enterprise varies widely. Having traditionally focused on the consumer segment, some MNOs are challenged in addressing the enterprise market which is where we believe the majority of value from edge computing services will come from. This is a critical factor in assessing the market potential for MNOs.

[1] End of 3Q2017.

# The Wireless Telco Landscape

## The Edge Computing Opportunity

Edge computing provides the following three opportunities to MNOs

1. Opex savings: primarily by reducing backhaul capacity between the edge and the core network. The business case for this application has not proven sufficiently attractive to MNOs based on the current track record for deploying mobile edge compute applications, especially before 2016. These edge applications focused on video caching. When OTTs encrypted video, edge caching became ineffective. Moreover, operators did not value the benefits of video edge caching and optimization. This could be attributed to the fact that LTE networks were still relatively unloaded with overhead capacity generally available across the network. Because demand on capacity is bounded by geography and time, operators could not rationalize the addition of hardware-based solutions to edge caching.
2. New services to the consumer segment: Edge computing brings benefits in latency that enables applications such as gaming, virtual and augmented reality, in addition to other benefits such as improving the quality of video. The challenge MNOs face is that consumers refrain from buying additional services and expect all benefits to be bundled in the subscription fee they already pay. Edge computing has to bring measurable benefits that consumers are willing to pay for. The ability of the MNO to monetize consumer services will depend on business models with the application or service providers and could include revenue sharing schemes.
3. New services to the enterprise segment: This represents the largest opportunity for MNOs and the one on which to base the edge compute business case. IoT applications in industry are an example. However, the MNOs will have competition from the Cloud players who are already engaged in building edge services and have a head start in the enterprise space over the MNOs. MNOs have not been traditionally successful in competing with OTTs. For example, Verizon exited the data center space, selling its 29 data centers to Equinix in December 2016 for $36 Billion.

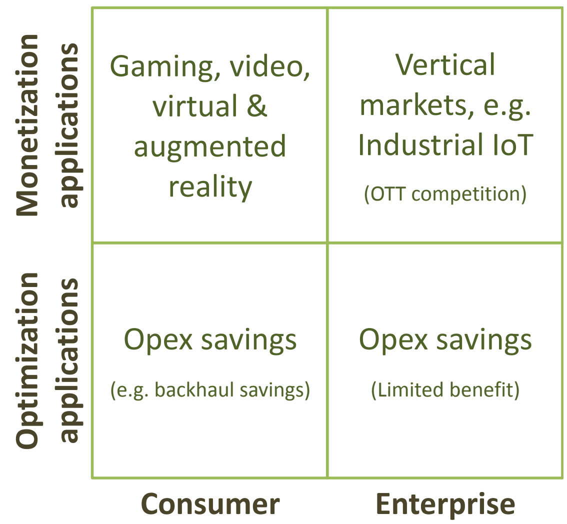



**Figure 2 Business opportunity for MNOs in edge computing.**

Understanding the opportunity edge computing brings together with the target market and applications is coupled with deployment scenarios for edge computing which we review later in this report.

## Network Status

Mobile network operators are in the midst of a few critical initiatives that potentially have an impact on deploying edge computing:

1- Driving virtualization into the network, focusing on the core network and OSS/BSS systems. This transformation is proceeding at different pace depending on the operator. Specifically, we find it common for operators to deploy virtualized solutions to support their IoT services.
2- Evaluating and planning for cloud RAN topology which centralizes baseband function into local edge data centers.
3- Evaluating and planning for 5G networks. Specifically, this includes features such as network slicing, in addition to edge computing.

Network virtualization is an enabler to edge computing. Aside from allowing applications to run on common hardware which greatly facilitates the practical implementation of edge computing, virtualization changes the cost structure, allowing MNOs to decentralize the core network. Many operators, especially in large countries such as AT&T and Verizon in the United States, opted to highly centralize their LTE core network into a few centers, not exceeding a dozen regional centers, to achieve economies of scale. Core virtualization allows these operators to distribute the core to a larger number of locations, bringing it closer to the user. In fact, AT&T indicated their plan to push the core to around 200 regional centers. While the exact number and plan is still being defined, we believe it is inevitable and a natural progress. This progress will favorably impact the implementation of edge computing in wireless networks.

The 5G architecture provides provisions to split the RAN into different functions that could be placed at different physical locations. This is meant to achieve a flexible architecture to meet the requirements of different applications and optimize the corresponding cost structure. Cloud RAN features data centers hosting centralized baseband processing which could be virtualized in software to run over commercial servers. Operators in Japan and Korea already implement a centralized RAN architecture. In the United States, Verizon is actively deploying fiber for transport service while pulling the baseband function from the cell site into baseband centers. This provides another synergy with edge computing which can reside in these baseband centers.

Geographically, the following conditions exist to implement edge compute:

a. United States: provision for edge compute by distributing the core network and leveraging central offices (AT&T) or implementing Cloud RAN (centralized baseband initially: Verizon).
b. Japan and Korea: Option to use central offices and centralized baseband centers (KT, STK, NTT Docomo).
c. China: Leverage central offices (China Telecom) and future CRAN locations (China Mobile).
d. Europe: Leverage central offices (Deutsche Telekom, Orange, Telefonica, BT EE, Telecom Italia).

An important takeaway related to virtualization: MNOs indicated to us their unwillingness to deploy proprietary and/or hardware-based edge solutions. Any edge computing solution needs to be based on a virtualized network architecture.

## MNO Outlook on Edge Computing

While virtualization is slowly sweeping through wireless networks operating LTE, operators we spoke with have been thinking of edge computing in terms of 5G networks. Edge computing could be implemented with 4G and 5G networks alike. However, the fact that operators put edge computing in context with 5G networks and features like network slicing indicates that timelines with edge computing will correlate more closely with 5G network uptake.

## MNO Edge Computing Activities

MNOs have been engaged in several market and technology development activities related to edge compute. Analysis of these engagements reveals that MNOs are still at the early stages of thinking through their edge compute strategy. We can confirm this through different conversations where MNOs are still in process of gauging the market opportunity and matching that with the edge compute architecture to validate the business case.

**Table 1 Summary of MNO edge compute activities.**

| Operator | Activities |
|---|---|
| **AT&T** | • Member of M-CORD, ETSI MEC<br>• Launched an edge computing test zone in its Palo Alto foundry to test applications such as connected cars, AR/VR and drones.<br>• *Foresee edge computing as part of 5G architecture, still in process of figuring out the business model* |
| **BT EE** | • Targeting MEC applications in critical communications and fixed wireless access,<br>• Pilot of MEC application for drones with Saguna and in stadiums with Nokia<br>• Participant in TIP MEC project |
| **China Mobile** | • Pilot projects with ZTE targeting IoT, VR and indoor positioning applications<br>• Pilot projects with Huawei targeting video content in stadiums leveraging small cells (LampSite) |
| **Deutsche Telekom** | • Collaboration with Carnegie Mellon on edge computing PoC and lab tests.<br>• Set up the hub:raum incubator hubs in Krakow, Poland and Berlin Germany to provide test bed for companies developing edge applications such as image recognition, telepresence, location and mapping, speech recognition, real-time drone control and augmented or virtual reality (2017)<br>• Set up a new edge computing business called MobiledgeX to explore opportunities in edge computing (2018). |
| **SK Telecom** | • Member of M-CORD<br>• Testing of edge computing as part of 5G network architecture in Samsung virtualized infrastructure. |
| **Telefonica** | • Participation in ETSI, TIP, and CORD |
| **Verizon** | • Member of M-CORD |
| **Vodafone** | • Member of Open Edge Computing and ETSI MEC<br>• Vodafone Group R&D collaboration with Saguna and Teragence to qualify edge compute impact on video traffic.<br>• Vodafone Australia and Nokia collaboration to build a proof of concept for edge computing applications in public safety, primarily video surveillance. |

## Edge Computing in 5G Networks

The objectives of 5G technology is to provide:

1. Ubiquitous device connectivity with uninterrupted user experience
2. Low latency (order of millisecond) for mission-critical systems and real-time applications, and enable services with zero-delay tolerance
3. High-speed gigabit connectivity

To achieve the objectives, 5G needs to provide a flexible architecture to meet the functional requirements (Table 2). A flexible architecture implies that network functions can be placed at tiered hierarchies. This contrasts with LTE, where radio functions are pushed to the base station. 5G networks will feature multiple hierarchies that are provisioned with the appropriate functions to serve specific use cases. Here, we summarize key technologies that will enable edge computing and improve its value proposition.

**Table 2 5G Functional Requirements.**

| Parameter | Value |
|---|---|
| **Air link latency** | < 1 msec |
| **End-to-end latency (device to core)** | < 10 msec |
| **Connection density** | 100x compared with LTE |
| **Area capacity density** | 1 Tbps/km$^2$ |
| **System spectral efficiency** | 10 bps/Hz/cell |
| **Peak throughput per connection (downlink)** | 10 Gbps |
| **Energy efficiency** | > 90% improvement over LTE |

Edge computing takes the form of an "Application Function" in the 5G network architecture which 3GPP defines as a function that could use the different capabilities of other network functions. Edge computing is one such function that can call upon the capabilities it requires. This culminates in a flexible framework that allows operators to deploy edge computing in different parts of the network: in the core, at the operator edge, or at the enterprise. This is a significant development because 4G networks did not make such provisions, centralizing many functions such as billing and legal intercept, that the edge computing paradigm broke. As a result, proprietary interfaces and fixed became necessary to implement edge computing in 4G networks which some operators were not in favor.

## Network Virtualization

Edge computing and NFV/SDN technologies are complementary and synergetic concepts. They reduce the barrier to entry for edge computing application developers and speed up the integration and service launch cycle. MNOs indicated their interest to implement virtualized edge compute applications on COTS hardware and to avoid proprietary, hardware-based solutions.

Implementation of virtualization in wireless networks is at a relatively early stage. Core network elements are the first candidates for virtualization: several MNOs have deployed virtualized OSS/BSS solutions and core network elements including vEPC, VoLTE gateways, PCRF and IMS nodes. Virtualization of the radio access network is technical complex and could lead to high fronthaul transport requirements. Nevertheless, the major infrastructure vendors are defining new RAN architectures leveraging virtualization. These architectures centralize the non-real-time layers of the air interface protocol stack while keep the real-time functions at the cell site. Leading operators such as AT&T, Verizon, SK Telecom, NTT Docomo and other reached advanced degree in implementing virtual core networks and are better positioned to reap the benefits from edge computing.

## Network Slicing

Network slicing allows provisioning instances or personalities of the network to serve applications with specific performance criteria. Network slicing leverages network virtualization concepts to create or remove network slices based on demand. While the full implementation of this technology is still a few years away, it integrates well with edge computing in that both technologies contribute to meeting the quality of service and experience subscribed to by the user.

While network slicing could be implemented on 4G networks, MNOs speak of network slicing in the context of 5G and virtualized networks.

## Private Enterprise Networks

The emergence of LTE in unlicensed spectrum (LAA, MuLTEfire) and shared spectrum regimes such as the 3.5 GHz Citizen Broadband Radio Service (CBRS) in the United States, enables enterprises to deploy private networks which can integrate edge computing. 5G builds on this background by further optimizing operation in different frequency bands: sub 6 GHz and millimeter waves, in licensed and unlicensed spectrum. The combination of virtualization, network slicing and unlicensed and shared spectrum regimes are catalysts for edge computing.

## Edge Computing Initiatives

A few variants and initiatives on edge computing exist under different names which we highlight through the prism of mobile networks. The variants have many elements in common, but with subtle differences due to their original backers.

### ETSI Multi-access Edge Compute

The ETSI Mobile Edge Compute ISG was formed in December 2014 to provide a standardized and open framework for edge compute platforms in mobile networks and to enable the integration of applications from vendors, service providers and third-parties (Figure 3). The scope of the ISG expanded beyond mobile access technologies to include other access technologies such as Wi-Fi and 5G access technologies such as fixed access in millimeter-wave bands. To reflect this new scope, the ISG was renamed in September 2016 to Multi-access Edge Compute (MEC). ETSI MEC released its first API specifications in July 2017.

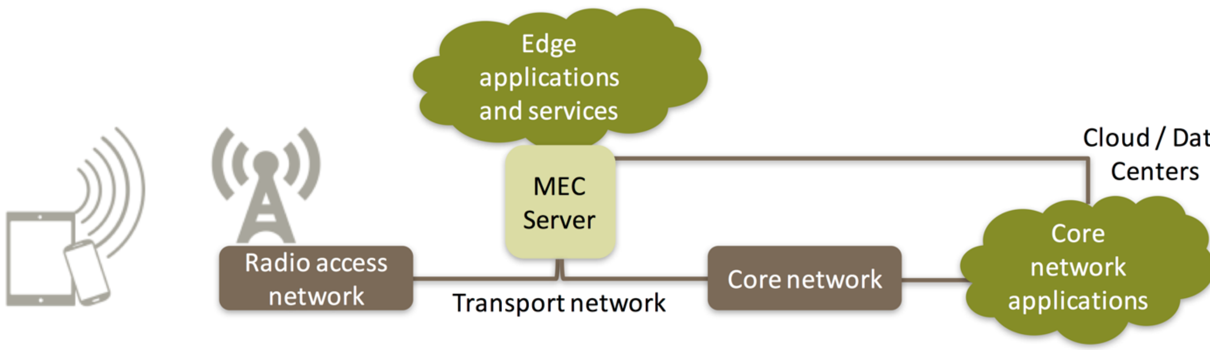



**Figure 3 MEC architecture overview.**

The roots of ETSI MEC are based in Nokia's Liquid Applications initiative, and Nokia remains a key player at the ETSI MEC. Other members include Huawei and Intel[2] in addition to major service providers, IT companies, and specialists developing varied MEC applications. Many leading players have dissented, most notably Ericsson, Verizon and Deutsche Telekom (Table 3).

**Table 3 MEC membership highlights.**

| Key Participants | Notable Absentees |
|---|---|
| TEMs: Nokia, Huawei, ZTE<br>Service providers: Vodafone, NTT Docomo, AT&T, Telefonica, Telecom Italia, Orange<br>Silicon: Intel, Xilinx<br>IT firms: IBM<br>Specialists: Saguna, PeerApp, Vasona, Argela, Quortus, ACS | TEMs: Ericsson<br>Service Providers: Deutsche Telekom, Verizon |

[2] Full list of members is available at https://portal.etsi.org/tbsitemap/mec/listofmembers.aspx

A technical overview of ETSI MEC is in Appendix 2. What is important to note is that the MEC architecture is based on NFV concepts. NFV is a critical element to edge computing. MNOs we spoke to indicated their interest to deploy NFV-based solutions only as proliferation of 'boxes' is no longer manageable.

ETSI MEC defined a number of proof of concept activities to demonstrate the use cases from which technical requirements were derived (Table 4).

**Table 4 ETSI MEC Proof-of-concept.**

| # | Description | MNO Lead | Ecosystem Players |
|---|---|---|---|
| PoC 1 | Video user experience optimization | China Mobile | Intel, iQiYi |
| PoC 2 | Edge Video Orchestration and Video Clip Replay | EE | Nokia, Smart Mobile Labs |
| PoC 3 | Radio aware video optimization in a fully virtualized network | Telecom Italia | Eurecom, Intel, Politecnico di Torino |
| PoC 4 | Flexible IP-based services | | City of Bristol, CVTC, University of Essex, Interdigital, Intracom |
| PoC 5 | Enterprise services | Bezeq | ADVA Optical Networking, Saguna |
| PoC 6 | Health care | Turk Telecom | Argela, Quortus |
| PoC 7 | Multiservice MEC platform for advanced service delivery | Vodafone | Advantech, Brocade, Cloudify, Saguna, Vasona |
| PoC 8 | Video analytics | Vodafone Australia | Nokia, SeeTec |

## OpenFog Consortium

Fog Computing is an edge compute concept introduced by Cisco in 2011 with a strong emphasis on IoT services. Fog Computing extends the edge to include the end-nodes or user devices. Consequently, Fog Computing Nodes (FCN) include storage, computing and communication capability to interact with other devices. FCNs provide awareness of device geolocation and device context, enable services such as translation between IP and non-IP transport in IoT applications, and decide on which network to join (e.g. 5G, LTE or Wi-Fi).

The OpenFog Consortium is a mixed public-private ecosystem founded in November 2015 to accelerate the adoption of Fog Computing in IoT applications, Artificial Intelligence, Robotics and the Tactile Internet. OpenFog counts 59 members including the founding members: ARM, Cisco, Dell, Intel, Microsoft and Princeton University Edge Computing Laboratory (notably, AWS is not a member of this organization). The OpenFog Consortium and the IEEE set up IEEE Project 1934 in November 2017 to develop standards for a computing and communication platform for interoperable information technology, communication technology and operation technology systems.

Proponents of Fog Computing differentiate it from Edge Computing by stressing that Fog is an extension of the cloud and works with it, whereas edge excludes the cloud. Fog addresses networking, storage, control and acceleration in addition to computing. Moreover, Fog is hierarchical whereas Edge includes a small number of layers.

The OpenFog Consortium and the ETSI MEC ISG have signed an MoU in September 2017 to cooperate on standards and interoperability requirements. They launched an initiative to adapt the MEC APIs for use in the OpenFog reference architecture to help developers create a common architecture, unify management and allow a single application software to run on both OpenFog and MEC architectures.

### Reference Architecture

The OpenFog reference architecture is an open, interoperable, horizontal system architecture for distributing computing, storage, control and networking functions closer to the users along a continuum of communication systems, computing nodes, and sensors.

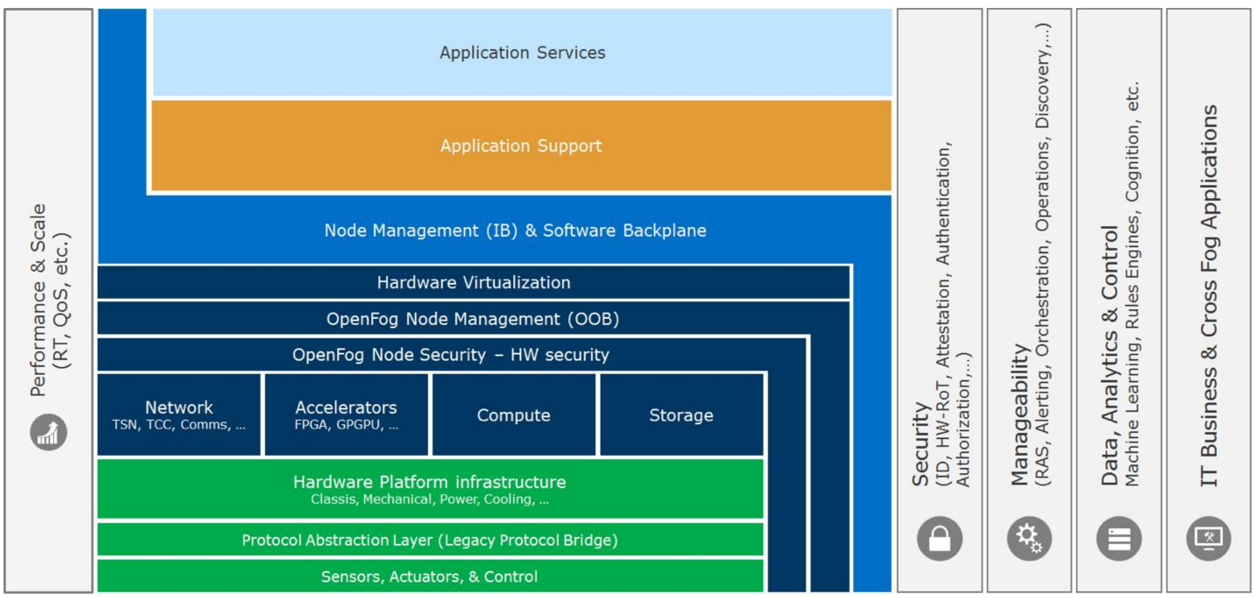



**Figure 4 OpenFog architecture description. [Source: OpenFog Consortium]**

### Open Edge Compute

Cloudlet denotes a small-scale mobile-enabled cloud data center that is close to the edge. It is the middle tier of a three-tier hierarchy: mobile device, cloudlet and cloud. Cloudlets are often associated with academia, primarily due to the leading work at Carnegie Mellon University (2013) in providing open source Cloudlet implementation mechanisms as extensions of OpenStack. The Microsoft micro-data centers as an extension of Azure hyper-scale Cloud data centers are also a representation of Cloudlets.

The Open Edge Computing (OEC) initiative is born out of the overlapping interest in Mobile Edge Compute and Cloudlets with an aim to leverage on-going research under the open source approach to accelerate market development. The Initiative builds on an Open Edge API and reference implementation that supports cloudlet-based deployments. Intel, Huawei, Vodafone, and Carnegie Mellon University launched the initiative in July 2015. A year later, Crown Castle, Nokia, Deutsche Telekom, NTT and Verizon joined the initiative. AT&T and British Telecom are industry partner candidates.

The Open Edge API is based on OpenStack implementation. It is the objectives of the initiative to make every OpenStack release edge-enabled. Open Edge also engaged with ETSI MEC ISG to contribute the Open Edge API in phase 2 of the ETSI MEC specifications. The Open Edge APIs map onto the Mx1 and Mx2 ETSI MEC APIs to third parties. OEC plans to launch the end-to-end Edge Platform with Open Edge API in 2018.

### Telecom Infra Project (TIP)

The Facebook-led TIP established the Edge Computing Working Group to focus on lab and field implementation for services and applications at the mobile and fixed network edge. The initiative will leverage existing architectures (e.g. ETSI MEC, M-CORD, vendor specific), libraries, software stacks and APIs instead of developing ones. The Group will develop design specifications for use cases proposed by its members and sponsored by a service provider or an enterprise who will lead a field trial implementation. The current use cases include the following:

1. Virtual CDN at the edge: Objective to deliver high-hit, low-latency content from the network edge. Project collaborators include Telefonica, Intel, Aricent, B.Yond (Nexius), and Quilt.
2. AR/VR Tourism: Objective to enable scene recognition activity using AR and develop new revenue service. Project collaborators include Deutsche Telekom, B.Yond (Nexius), Intel, Telefonica and Quilt.
3. Drone video delivery: Objective to get real-time video from drones for events, venues and attractions. Project collaborators include Deutsche Telekom, Telefonica and Bai Cells.

### Mobile Central Office Re-architected as a Data Center (M-CORD)

M-CORD is a joint project between Stanford University's ON.Lab and The Linux Foundation, with participation by AT&T, SK Telecom, Verizon and NTT and other ecosystem players (silicon, software stack, test equipment, etc.). M-CORD is defining a blueprint for virtualized mobile network architecture and looks to transform central offices into data centers. Activities at M-CORD include disaggregation of the control and user planes of the radio access network and packet core, network slicing and edge computing.

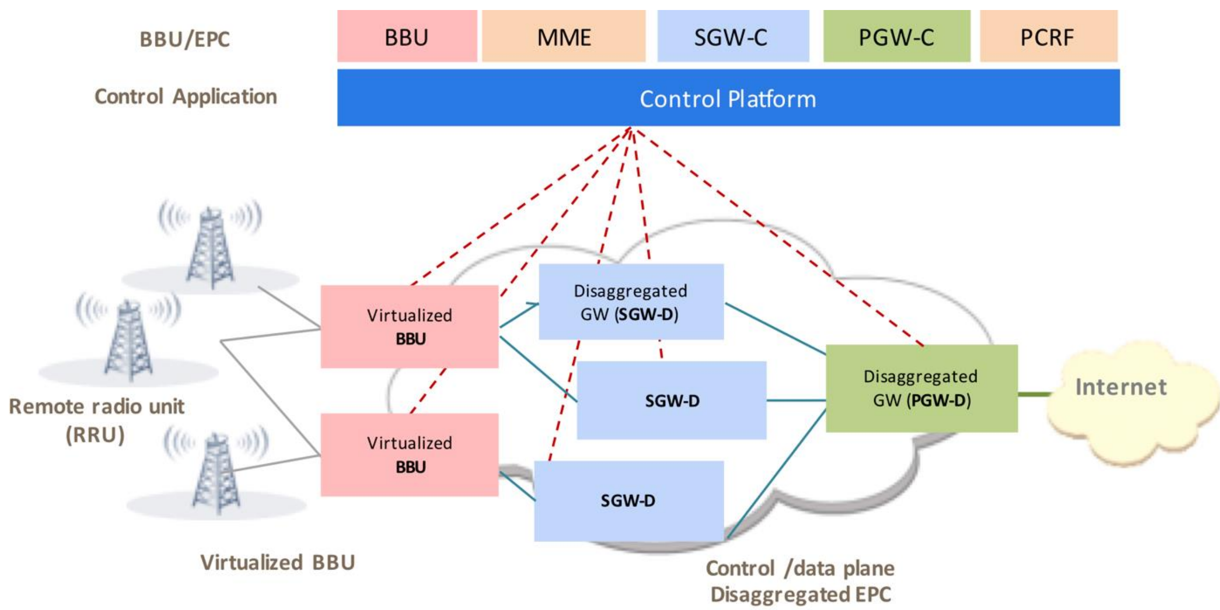



**Figure 5 M-CORD Reference Architecture. [Source: M-CORD]**

The idea of turning central offices into data centers is powerful and matches with the objectives of edge computing. Mobile network operators have significant number of central offices at their disposal, especially the MNOs who operated a POTS network before the advent of the mobile technology. MNOs without fixed access assets have multiple switching offices for hosting 3G RNCs or 2G BSCs which could be utilized as locations to host edge computing. To illustrate, AT&T has 4700 central offices compared to only 11 national centers where it hosts the LTE packet core. China Mobile operates over 400 central offices in Shanghai alone. In Europe, Deutsche Telekom still possess over 900 central offices across Germany even as it sold many other locations.

## Observations on Edge Computing Initiatives

We expect for the ETSI MEC to increase collaboration with other initiatives and to continue to broaden its architecture which remains to date focused on the wireless network. However, connectivity is not restricted to 3GPP-based technologies and many other technologies exist particularly in an enterprise environment.

Many leading MNOs are not amenable to the initial ETSI MEC architecture which places edge computing server between the core network and the base station. They cite that the architecture breaks the 3GPP framework for billing and legal intercept. Instead, they are looking at placing the edge compute platform behind the EPC while pushing the EPC further down towards the edge. We expect this architecture to dominate edge computing deployments in the mobile network.

## Edge Computing Applications

The financial benefits of edge computing can be divided into two categories: monetization applications that contribute to revenue generation, and applications that lead to operational cost savings (Table 5). While often the two categories are related, it helps to think of applications in this framework to identify which contribute directly to revenue, as MNOs indicated higher interest in this category.

**Table 5 MEC requirements categorized according to revenue and cost.**

| Requirements | Monetization Applications: Revenue Generation | Optimization Applications: Cost Savings |
|---|---|---|
| High responsiveness, low latency, near real-time operation | ✓ | |
| Data caching | | ✓ |
| Context-aware services | ✓ | |
| Location-aware services | ✓ | |
| Heavy computation applications | ✓ | ✓ |
| Data transformation and transcoding | | ✓ |
| Extended battery operation | ✓ | |

**Table 6 Optimization vs. Monetization Applications.**

| Optimization Applications | Monetization Applications |
|---|---|
| • Business case tilted towards reducing operating expenses<br>• Particularly useful for consumer segment<br>• Better economic viability with edge computing closer to core: amortizing expense over larger number of subscribers and base stations<br>• Amenable to deployments by MNOs | • Business case revolves around generating new revenues<br>• Particularly useful for enterprise segment<br>• Better performance advantage in placing edge computing closer to user<br>• Amenable to deployments by MNOs and third-party service providers |

Edge computing provides different benefits depending on the application (Table 7). It's important to note that some of the below applications can only be delivered with the implementation of edge computing services, as otherwise their requirements cannot be met by the wireless network. For example, some automotive and smart grids applications require sub 1 ms latency, which can only be achieved with edge computing[3].

---

[3] The 1 ms figure is frequently used in industry, but in reality, consensus converging on < 5 ms latency requirement.

**Table 7 Edge computing applications and their corresponding requirements.**

| Applications | Latency | Caching | Context | Location | Computation | Transcoding | Power | Deployment Timeline |
|---|---|---|---|---|---|---|---|---|
| **Content Caching** | ✓ | ✓ | | ✓ | | | | M/L |
| **Traffic optimization** | | | ✓ | ✓ | | ✓ | | S |
| **Augmented reality** | ✓ | | | | ✓ | | | L |
| **Virtual reality** | ✓ | | | | ✓ | | | L |
| **Multimedia content delivery (video)** | | ✓ | | | | ✓ | | S |
| *Enterprise applications* | | | | | | | | |
| **Asset tracking** | | | | ✓ | | | ✓ | M |
| **Video surveillance & analytics** | | | | | ✓ | ✓ | | M |
| **Local voice and data routing** | | ✓ | | ✓ | | | | M |
| *Retail services* | | | | | | | | |
| **Ad delivery** | | | ✓ | ✓ | | | | M |
| **Footprint analysis** | | | ✓ | ✓ | | | | M |
| *IoT Connectivity* | | | | | | | | |
| **Massive IoT (e.g. sensor or meter reading)** | | | | | ✓ | | ✓ | M/L |
| **Critical IoT (e.g. smart grid switching, fault detection)** | ✓ | | | | | | | M/L |
| *Critical Communications* | | | | | | | | |
| **Traffic safety and control systems** | ✓ | | | | | | | L |
| **Precision farming** | ✓ | | | | | | | L |
| **Industrial IoT, time critical process control** | ✓ | | | | | | | M/L |
| **Hazard warning** | ✓ | | | | | | | L |
| **Cooperative autonomous driving** | ✓ | | | | | | | L |
| **Healthcare applications** | ✓ | | ✓ | ✓ | | | | M/L |
| Deployment timeline: S: short term; < 3 years \| M: Medium term; 3-5 years \| L: Long term > 5 years | | | | | | | | |

The most discussed MEC applications to date include:

1- **Traffic optimization**: Traffic classification, steering, buffering and transcoding are techniques that may be used in conjunctions with policies to improve wireless network

performance: reduce streaming hang-ups and interruptions, and improve signaling round-trip time and retransmissions. Video is the driver for traffic optimization, as it constitutes 50% of today's mobile traffic and is rapidly growing through new applications like live streaming on social media. Video traffic is projected to grow to 75% of total mobile traffic in 2023[4]. Video streaming quality is increasingly becoming a parameter that leads to churn, as video application users face more issues that others: a study by Ericsson found that smartphone users who experience a high number of issues (> 11 per week) are twice as likely to contemplate switching service providers[5].
The deployment of traffic optimization is still at relatively early stage with a few leading operators involved in market trials and limited deployments. We expect traffic optimization engines to get higher attention from MNOs for dense urban areas and in markets where spectrum is relatively limited. However, in markets where spectrum is abundant, MNOs would opt to buildout additional infrastructure.

2- **Enterprise applications**: This encompasses a wide set of applications often in conjunction with small cell deployments. Two types of deployments emerge: private enterprises, where the network is used primarily by the enterprise employees, and public enterprise networks that are open to the service the public. An example of the latter is a stadium deployment with multimedia broadcast multicast services (eMBMS). The challenge in private networks centers on the ability to support multi-operator networks, which has dampened uptake in this market. This is set to change with new spectrum regimes such as CBRS shared spectrum, which we cover later in this report. Other challenges are related to the business model where all entities in the value chain – the enterprise, the edge computing service provider and network operator – need to benefit.

3- **Content caching**: Content caching reduces backhaul costs and improves response time for applications such as video streaming. This application did not have much commercial success: MNOs could not justify the cost of providing improved experience to their subscribers and when traffic became encrypted, MNOs could no longer implement it at all. This caused many caching solution providers to change business plans.

Today, the majority of Internet traffic is encrypted (>50%). Major OTTs and sites such as Facebook, Twitter, Google, Wikipedia, Wordpress and many others adopted the HTTPS protocol which encrypts data in transit. Netflix which is a major source of North American traffic began encrypting traffic for certain web browsers, although the majority of their traffic remains unencrypted. Encrypting traffic negates the opportunity of service providers to cache traffic, but for specific cases where they have control over

[4] Ericsson Mobility Report, November 2017.
[5] Ericsson Consumer Lab and Industry Insight Report: Experience shapes mobile customer loyalty, August 2016.

the traffic, such as in a closed-loop deployment in enterprises, stadiums and other venues.

## Latency Analysis

### Application Requirements

While it's been common to claim the need for sub 1 msec latency, most applications don't require such latency (Table 8). In fact, the industry is converging on a target latency of 5 msec instead of 1 msec and active ecosystem participants have shifted their messaging on this topic. This is in large part to the realization that a 1 msec latency will be expensive to achieve over public networks with undetermined and unproven opportunity for return on investment.

**Table 8 Latency Requirements**

| Latency | Applications |
|---|---|
| **≤ 1 msec** | • Telepresence with real-time synchronous haptic feedback<br>• Industrial robots<br>• Closed-loop industrial control systems<br>• Negotiated automatic cooperative driving<br>• Smart grid: Synchronous phasing of power supplies |
| **≤ 10 msec** | • Shared haptic virtual environment<br>• Tele-medical applications, e.g. tele-diagnosis, tele-rehabilitation<br>• Augmented reality<br>• Haptic overlay trainer/learner for fine motor skills (e.g. for medical)<br>• Smart grids<br>• Process automation<br>• Cooperative Collision Avoidance<br>• High-density platooning |
| **≤ 50 msec** | • Serious gaming (20 msec)<br>• Cognitive assistance (20-40 msec)<br>• Virtual reality<br>• Cooperative driving (20 msec)<br>• UAV control (10 – 50 msec)<br>• Remote robot control with haptic feedback (25 msec)<br>• Automotive pre-crash sensing warning |
| **≤ 100 msec** | • Vehicle safety apps (mutual awareness of vehicles for warning/alerting)<br>• Assisted driving – cars make cooperative decisions, but driver stays in control |

Applications play a critical part in qualifying the potential edge computing opportunity for MNOs. Some of the applications indicated here are nascent, while others exist today and are based on different technologies. For example, SCADA and WiSUN are used extensively in smart grids. Recognizing such alternatives and competition helps to moderate excessive forecasts of the market opportunity.

## Latency in Wireless Networks

The total latency in networks is due to the contribution of different elements in the network (Table 9). Field tests demonstrated that end-to-end latency varies widely when the application runs in the core, far removed from the edge (jitter). A key benefit for edge computing is to reduce the latency jitter to a narrow margin that's acceptable for the application and end customer. Field trials proved that edge computing reduces jitter and latency when the same application is moved to run closer to the edge.

**Table 9 Latency contribution factors.**

| Type | Description | How to reduce latency |
|---|---|---|
| **Air interface** | Delay due to frame structure and medium access control layer | Reduce frame duration, improve scheduler performance, improve medium access process.<br>LTE frame structure is 10 msec with 1 msec subframe. 5G maintains the same frame structure but improves the Transmission Time Interval (TTI) from 1 msec to as low as 0.14 msec to reduce air interface latency. |
| **Network element processing delay** | Delay due to the processing delay of network elements in the path between the user and the data servers. These elements include the base station, transport routers or switches, core network elements such as S-GW and P-GW, data or video servers, etc. Network elements introduce jitter that leads to relatively wide variations in latency around an average value. | Depending on the type of the element, the latency could be very small – on the order of micro seconds – which means there's little benefit from reducing it. Alternatively, it could be on the order of a few milliseconds in routers. Edge computing reduces the number of network elements between the application and the edge which reduces both latency and jitter. |
| **Transport distance** | Delay due to the physical separation between the user and data source. | Moving the data source, or more generally computing and storage closer to the user which is what edge compute calls for. |

For comparative analysis, the average round-trip latency (RTT) on LTE networks globally is 48 msec. A few small networks, such as Swisscom, reached RTT of 27 msec latency. In the US, the average LTE RTT is 66 msec. To further reduce end-to-end latency, 5G will provide for short frame duration and an optimized medium access control and scheduling functions to reduce latency. Moreover, it is possible to achieve lower latency on private networks than public carrier networks. Private networks could be designed to better meet user specifications than public networks. This is an important angle to consider in the overall market for edge compute in wireless networks, particularly in the enterprise.

## Edge Computing Deployments

The location of edge computing is a tradeoff between performance and cost. It is possible to assess the viability of the deployment location from the perspective of the main business benefit:

1. **Latency benefit – monetization driver**: the closer edge computing is placed to the edge, the lower the latency. This is an important element in meeting the requirements of 5G target of 1 ms latency: edge compute is a prerequisite to meet 5G network requirements. The latency benefit enables service providers to offer and monetize new services, particularly enterprise services where there is a higher willingness to pay.

2. **Backhaul benefit – optimization driver**: edge computing reduces backhaul bandwidth, as well as signaling load towards the core network and the Internet. Consequently, it results in financial savings to service providers which varies depending on the backhaul network architecture. Typical savings are benchmarked at up to 35%.

It is possible to deploy edge computing in different locations (Figure 6):

1. **Option 1: At the base station** – this is an extreme case that provides the best performance at the highest cost.
2. **Option 2: At an enterprise site** – this is applicable to a small cell network in an enterprise. The edge applications are optimized to support services desired by the enterprise. There is a business risk in that both the service provider and the enterprise need to validate the business opportunity and return on investment.
3. **Option 3: At a hub site** – this is where edge compute serves a number of base stations from a backhaul aggregation site (typical macro cells). The hub site could also be the Cloud RAN baseband center where all the base station basebands modules are collocated.
4. **Option 4: Edge data center** – this case features a core network that is placed in an edge data center such as a central office serving a city. The data center could include an instantiation of the virtual core network. This model balances cost efficiency with an architecture that's compatible with 3GPP LTE topology. It is also compatible with 5G network architecture.

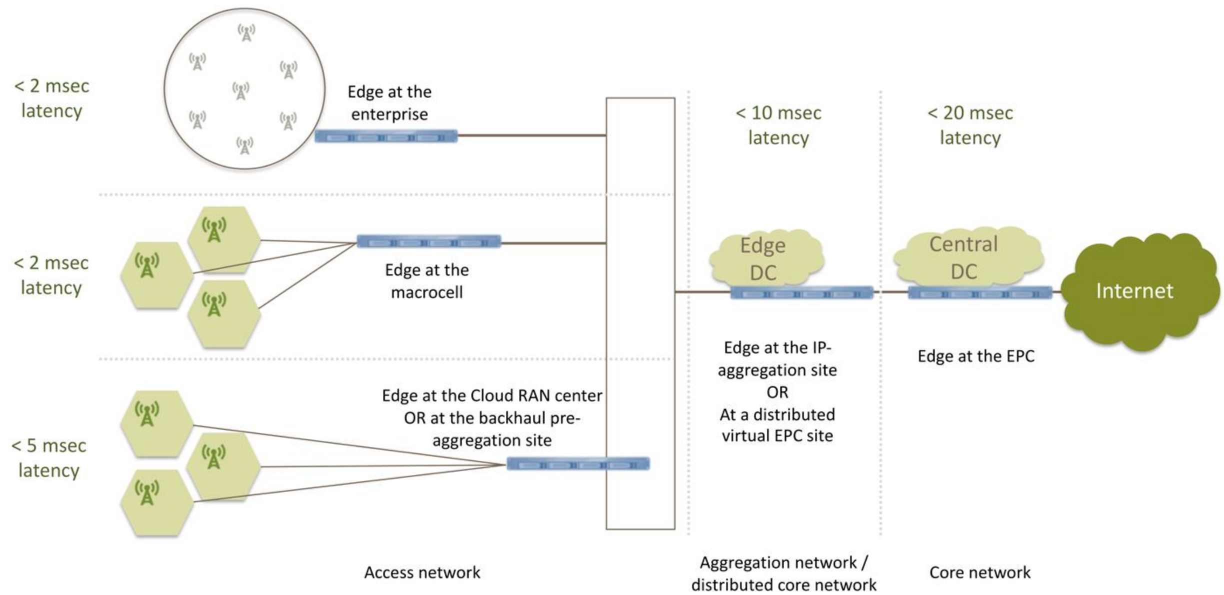



**Figure 6 Edge computing deployment scenarios.**

While there is currently no consensus on where edge computing will be deployed in the network due to varying requirements, there are two clear trends emerging, and one additional trend that's yet to prove its practicality:

1. **Operators to exclude edge computing from the macro cell site**: this is mainly due to the high cost of deployment. Operators are concerned that they will not be able to achieve positive RoI with this architecture.
2. **Operators strongly considering edge data centers**. This is combined with the drive to virtualize the core network and distribute the new virtual core towards edge data centers. Operators like AT&T, Verizon and other leading operators are moving along this path which is contingent on core network virtualization. We expect this path to be the dominant approach for network operators as it balances the cost of the deployment against potential business case. This model allows MNOs to compete with OTTs.
3. **Edge at the enterprise**: this model has a number of technical, financial, operational and regulatory dependencies impacting its evolution. Implementing edge at the enterprise level implies the use of small cells. This could be achieved with private networks where the following challenges are present:
    a. Core network scalability: the traditional core network is designed for scale to support millions of subscribers. Scalability to support a relatively small number of devices is poor resulting in high cost per connection. Virtualization is set to change this. The business models are still in the early stage of development as vendors find it challenging to monetize small deployments.
    b. Spectrum: Service providers control spectrum assets. Private network operators need to deploy through a service provider. The use of shared spectrum (e.g. CBRS) and unlicensed spectrum (e.g. 5 GHz) alleviates this issue.
    c. Positive ROI for the enterprise and the service provider has been slow to develop.

d. Many service providers lack the business channels into the enterprise as they focus on the consumer segment. Service providers would have to transform from a B2C model to B2B or B2B2C model.

## Edge Computing Business Case

The business case for edge computing has high variance related to both cost and revenue sides of the equation. The location of the edge computing hardware determines both the cost and benefits – it is a tradeoff between performance and cost (Figure 7). To address this, the architects of 5G networks designed support for multiple hierarchies, unlike the centralized architecture of 4G.

Virtualization of the mobile network will positively impact the business case for edge computing. Applications developed in software running over commercial hardware speeds up service introduction and improves operational flexibility.

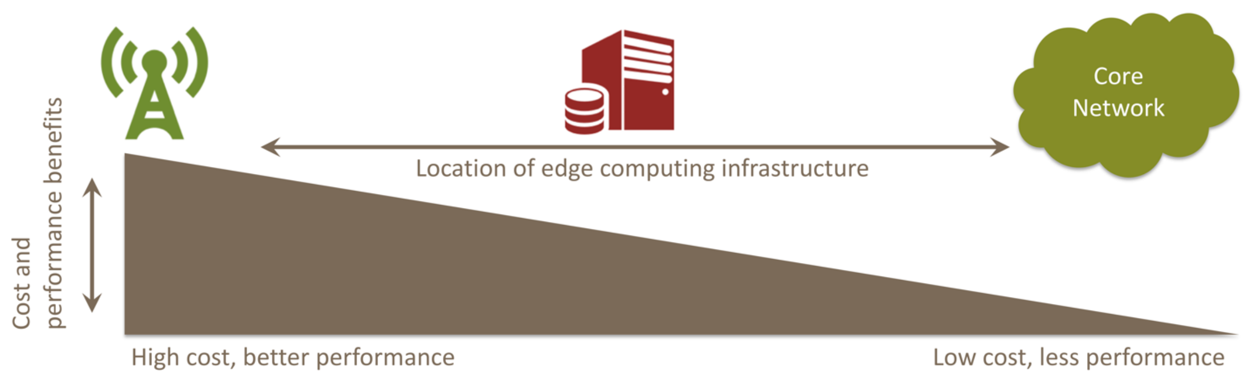



**Figure 7 Edge computing infrastructure location vs. performance and cost tradeoff.**

## Edge Computing Business Models

In a classical mode, the MNO deploys edge computing to monetize services they provide directly to the enterprise or to consumers. But other business models exist for edge computing services, including those provided by neutral hosts or Mobile Virtual Network Operators (MVNOs). These models are yet to emerge and could serve to complement MNOs services.

### Neutral Hosts Business Model

Having an independent entity provide differentiated services creates new monetization opportunities. Mobile network operators base their business on serving the consumer segment, but lag in providing services to enterprises. Enterprises are OT-focused with efforts to integrate IT to improve efficiency and performance. MNOs focus on telecom services, primarily connectivity. Serving the enterprise effectively requires MNOs to approach the enterprise from an IT perspective, at a minimum, without forgetting the importance of OT. Third-parties who possess both IT and OT competencies would have more success to serving the fragmented enterprise market than MNOs going direct. Neutral hosts would be capable of tailoring services to their clients' requirements and more quickly responding to market needs than MNOs.

Success of third-party edge computing service providers would depend on the MNO's willingness to engage, and the extent to which interfaces will be open and made available for neutral hosts. An attractive business case will compel the MNOs to engage. This will subsequently drive service integration by opening the appropriate interfaces. Interoperability of edge computing platforms with the wireless network infrastructure is therefore a critical issue.

How third-party entities will engage is still to emerge. Private networks, based on small cells and virtual RAN solutions operating in unlicensed and shared spectrum bands represent the opportunity for this model to take off. We anticipate this model would be tested in the United States first, in the 3.5 GHz CBRS band, followed later in other parts of the world in unlicensed spectrum (5 GHz). Use of these spectrum bands limits interference with service provider's macro-cellular network, while indoor deployments limit adverse mobility management impact. Thus, private virtual RAN or small cell networks reduce the risk related to network planning and deployment by third parties, which is a major concern to network operators. Based on this, we anticipate that edge computing will receive a boost from the advent of shared spectrum and the implementation of cellular technologies in unlicensed spectrum.

The neutral host business model could be coupled with the Mobile Virtual Network Operator business model to provide a comprehensive service offering to enterprises. This business model frees network operators to focus on different services and markets according to their business strategy.

**Table 10 SWOT analysis for edge computing operation by third parties.**

| Strengths | Weakness |
|---|---|
| • Address enterprise market with value added services to generate new revenue opportunities<br>• Insightful understanding of enterprise requirements<br>• Integrates applications and delivers on enterprise requirements that are underserved by network operators | • Emerging market in the process of developing the value proposition<br>• Reliance on small cells and private networks limits the scope of this model in some deployment scenarios |
| **Opportunities** | **Threats** |
| • Many enterprise applications exist and are yet unfulfilled<br>• Shared and unlicensed spectrum regimes propel private networks and services by third parties | • Requires open interfaces for seamless integration into the mobile infrastructure network<br>• Network operators' willingness to open their infrastructure will depend on the business case as well as competitive and strategic factors |

### Edge Computing Services by Network Operators

The impetus for edge computing deployments by service providers is twofold. First are operational savings resulting from a reduction in backhaul traffic. This is a relatively small opportunity from a revenue and market perspective. Second, is monetization services resulting from new applications which represent the bulk of the value that edge computing presents.

Edge computing services by MNOs is applicable to both the consumer and enterprise market segments. It has the potential to change how content providers and OTT players interact with network operators. In both segments, MNOs would have the opportunity to launch new services to claim some of the value that OTTs derive through their services. To capitalize on this opportunity, MNOs will have to strike partnerships with application and content providers as well as devise service valued by both the consumer and enterprise market segments. Mobile accounts for much of the OTTs revenues – e.g. more than 1.2 billion Facebook and 700 million YouTube users. Similarly, the Internet giants are actively seeking great share of enterprise services leveraging their Cloud and edge Cloud infrastructure. Edge computing is a tool for the MNOs to compete with the OTTs and claim value in addition to commoditized connectivity services.

**Table 11 SWOT analysis for edge computing operation by network operators.**

| Strengths | Weakness |
|---|---|
| • Integration with the mobile infrastructure<br>• Highly applicable to target the consumer segment by network operators<br>• Single entity for service delivery, operations and management | • Need to develop better understanding for enterprise requirements and adapt culture and operation to serve the enterprise sector<br>• Address the consumer segment with services they are willing to pay for |
| **Opportunities** | **Threats** |
| • Allows network operators to develop and monetize new services<br>• Allows network operators to derive operational savings<br>• Emerging opportunities in vertical markets for enterprise services | • Raises questions related to net-neutrality<br>• Edge computing is an IT, not a RAN solution, which offers a different paradigm than established norms<br>• Edge computing has different requirements and breaks down standard operating processes within network operators leading to inter-organizational resistance<br>• Competition from OTTs in both the consumer and enterprise segments |

## Challenges to Edge Computing Implementation and Monetization

Service providers must overcome numerous challenges in order to capitalize on the opportunities edge computing provides. Among the technological challenges, we note:

- Integration with 4G mobile networks. This is a result of misalignment between edge computing and LTE network architecture which centralizes functions such as billing and legal intercept at the core network. Edge computing fractures that architecture as data flow concentrates at the edge and does not pass through the core. In order to address this, edge computing solutions implement proprietary interfaces. MNOs have been hesitant to implement proprietary solutions. This inhibits local breakout solutions from taking firm hold in the market. 5G promise to resolve this issue through a flexible architecture that leverages virtualization. This makes the growth of edge computing tied to a large degree to the proliferation of 5G networks and the extent of network virtualization.
- Security and network integrity concerns by network operators, as edge computing provides an open environment for third party applications to run on the telecom infrastructure. The success of edge computing is in part reliant on participation by third party application developers and service providers. MNOs will likely want to maintain tight control over the applications running on the edge computing server. Therefore, security concerns by MNOs are likely to impede the growth of edge computing.

- Maintaining service over several radio access technologies that characterize heterogeneous networks, such as LTE, Wi-Fi and 5G technologies. This issue will be addressed as through the architecture of 5G and specifications by the different edge computing initiatives such as ETSI MEC, but it does add a layer of complexity to achieve full integration across multiple types of networks.
- Spectrum in the case where edge computing in the enterprise segment leverages deployment of small cells. MNOs have been hesitant to deploy small cells for a number of reasons including interference with the macrocell layer. Availability of spectrum for small cell deployments, such as the CBRS band in the United States, will favorably impact the implementation of edge computing.

From a commercial perspective, the challenges include:

- Validating the business case for an open edge computing environment to both service providers and the beneficiaries of edge computing in both enterprises and consumer segments. The business case varies depending on the application and it is often hard to quantify, especially for future applications such as VR/AR and V2X services. Validating such business case requires active interaction with vertical market players, which is starting to happen.
- Defining a successful business model for edge computing services by third parties, where MNOs would find a compelling reason to actively participate in making the business a success. We expect that revenue sharing models to emerge between MNOs and such service providers to capitalize on the edge computing architecture.
- Competition with the Internet giants such as AWS and Microsoft and other OTTs who are in process of defining and offering edge computing services.
- Consumer willingness to pay for edge computing services such as enhanced video, online gaming, and other applications.
- Handling of content including digital rights and content access management, encryption and storage of the content within the network.
- Issues related to net neutrality where OTTs would have similar access to network resources as service providers.

Finally, culture and business operation are important challenge from both supply and demand side of edge computing services. On the supply side, the telecom service providers approach edge computing from a telecom not an IT perspective. On the demand side, enterprise operation technology (OT) is a key function that at times clashes with information technology. This acts to slow down the edge computing from a demand side.

## Return on Investment Analysis

To illustrate the optimal scenario for deployment of Edge Computing, we investigated a hypothetical case study, for an operator with both enterprise and consumer customers.

The case study assumes:

- A network with 100 million subscribers
- 100,000 radio site locations
- 200 regional data centers
- 5 centralized "original core network" sites
- $8,000 per radio sector (3 sectors per site)
- Cheaper standard computing at data centers equal to $6 per subscriber
- Backhaul cost of $5 billion to cover all 100K cell sites
- Enterprise applications:
  - 1000 enterprise locations willing to pay $5K per month for low latency connectivity
- Consumer applications:
  - 5% of subscribers are willing to pay $5 per month for low latency gaming
  - 10% of subscribers are willing to pay $8 per month for locally tailored video

Given those assumptions, Mobile Experts calculated two key factors:

1. The savings in deployment of backhaul (fiber) to support network traffic. We estimate that 15% of fiber costs can be saved with edge computing at radio sites.
2. The revenue opportunities from locally hosted services.

Both of these two economic incentives are important, as the cost of deploying computing platforms on 100,000 sites can be daunting. In the enterprise case, we find that some ultra-low latency applications such as controlled robotics can only be managed through local edge computing, so the revenue opportunity in these cases is only available in the full EC scenario.

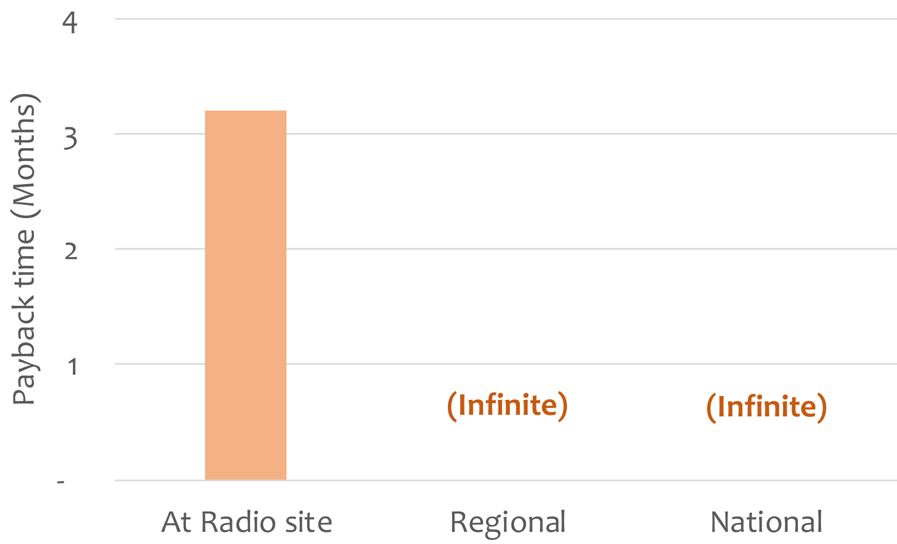



**Figure 8 Return on Investment (payback time) for enterprise robotics applications**

In the consumer "gaming" scenario, the optimal investment is to have EC in the regional data centers, with a compromise between latency for the user and CAPEX. This preliminary conclusion could change over time, as gaming or other applications emerge in which low latency is critical. In today's environment, where regional data centers can offer an adequate but not ideal user experience, even if only half of the gamers would adopt the system the ROI would be better than wider adoption.

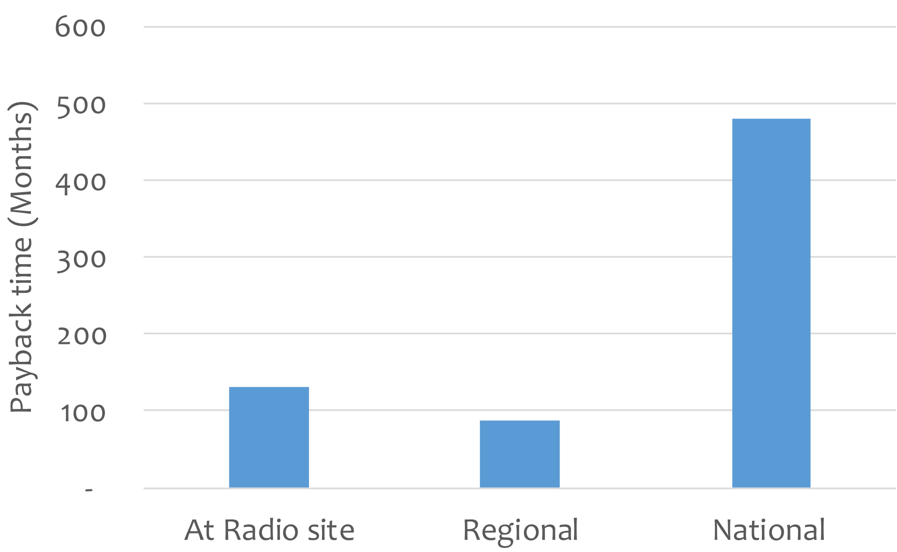



**Figure 9 Return on Investment (payback time) for gaming applications**

Finally, in a consumer case for video services that are tailored for local consumers, (and latency is not as important as the backhaul savings), a compromise at the regional data center appears to be the best fit.   .

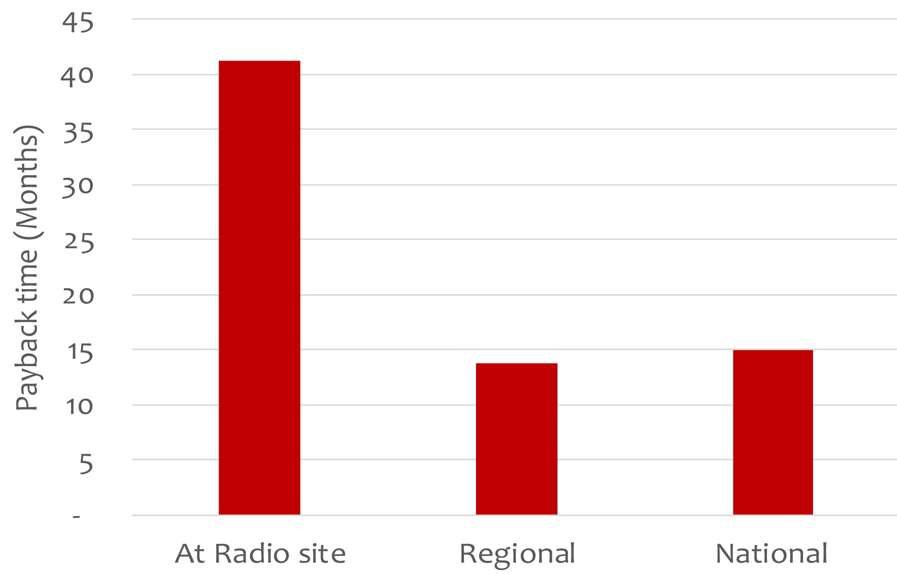



**Figure 10 Return on Investment (payback time) for local video applications**

Each of the above scenarios assumes that a single application must carry the burden of paying for the edge computing resources.    If we combine all three applications in a more holistic view of a business case, again we see that the regional data center is the best choice.

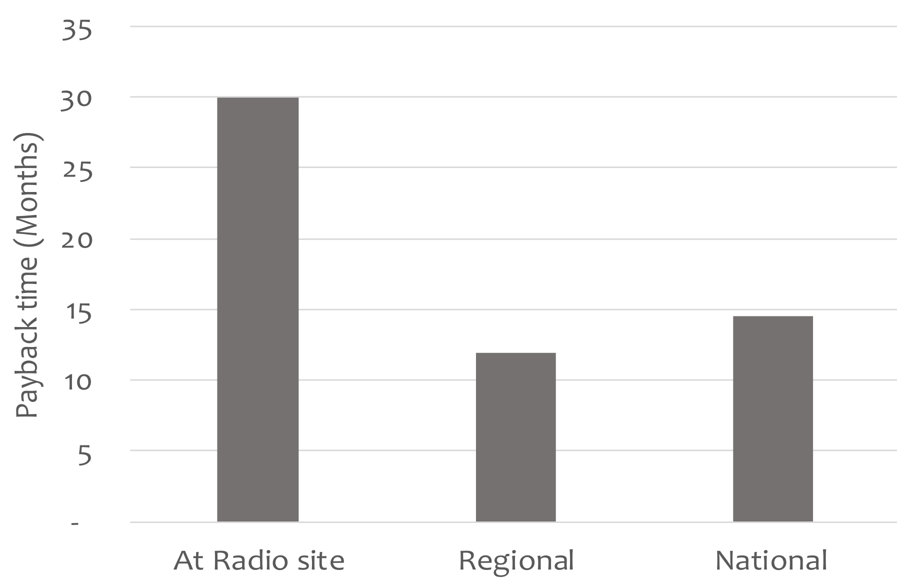



**Figure 11 ROI for a combination of enterprise and consumer applications**

## Edge Compute Players

***Telecom and Networking Equipment Manufacturers***
Adva
Altiostar
Cisco
Ericsson
Fujitsu
Huawei
Juniper
Motorola Mobility
NEC
Nokia
Samsung
ZTE

***Edge computing specialists (platforms, applications, protocol stacks)***
Allot
Netsia
PeerApp
Qwilt
Saguna
Vasona

***Gateways and Core Network Elements***
Astri
Athonet
Brocade
Quartus

***Small Cell Players***
Accelleran
BaiCells

***EC Application Developers & Integrators***
ACS
Openet
PoLTE

***Silicon Vendors***
AMD
Cavium
Intel
Nvidia
Xilinx

***Servers and EC Infrastructure***
Adlink
Advantech
Artesyn

***OS and Software Middleware***
Gigaspaces
RedHat

***System Integration and NFV***
HPE
IBM
Italtel
Tech Mahindra

Additionally, many research organizations and academic institutions are involved in edge computing.

## Edge Computing in the Field

The following are a few examples of edge computing field activities for illustrative purpose:

**Nokia:** Nokia is very active in edge computing activities and has initiated and concluded a number of trials including:

- A deployment with China Mobile at the Shanghai F1 that included MEC with small cells. The system was used to decrease the latency for video streaming within the venue.
- Field trial with Google of MEC impact on video performance.
- Construction of a test bed in collaboration with Deutsch Telekom, Continental and Fraunhofer ESK to exchange traffic warning messages to vehicles in under 20 msec.

**Huawei**: Together with China Mobile, Huawei coupled edge compute applications with the Service Anchor network solution and LampSite small cells deployed in a shopping mall to provide indoor location-based services including indoor navigation, vehicle location, mobile sign-in, human traffic flow identification, and targeted marketing.

**PeerApp**: Currently running multiple market trials of its traffic optimization and caching solution with MNOs in Asia Pacific, Africa, and Latin America.

**Vasona**: Ongoing market deployment of Vasona's traffic optimization solution by Telefonica UK and an American service provider. At Mobile World Congress 2018, Vasona demonstrated an online gaming edge computing application that eliminates the latency bottle neck to enable gaming services that otherwise could not be supported by the network, thereby providing a clear example of how service providers could monetize consumer applications, which in this case would involve revenue sharing model with an online gaming service provider.

**Saguna**: Initially focused on caching solutions, which it trialed at a few MNOs including some in partnership with Nokia. Saguna shifted from caching to focus on developing edge computing platforms. In collaboration with Vodafone, Saguna demonstrated and quantified the benefits of edge computing in improving video services.

**Juniper**: Collaborated with Saguna to run VNFs on Saguna's Open-RAN MEC platform.

**Netsia and Quortus**: Demonstrated a proof of concept of an edge computing solution for the healthcare vertical in collaboration with Turk Telecom. The demonstration included network slicing capabilities using Netsia's ProgRAN RAN slicing solution running on small cells that are connected to a Quortus vEPC solution.

Observations from field trials lead us to conclude that emphasis is shifting towards monetization applications, although some optimization applications such as traffic optimization are still promising.

## Market Forecast

We project edge computing to generate a cumulative revenue of $7.2 billion, through 2023, in both the consumer and the enterprise markets. The enterprise market would dominate the revenue stream generating 61% of the cumulative revenue while the consumer segment generates the remaining 39%.

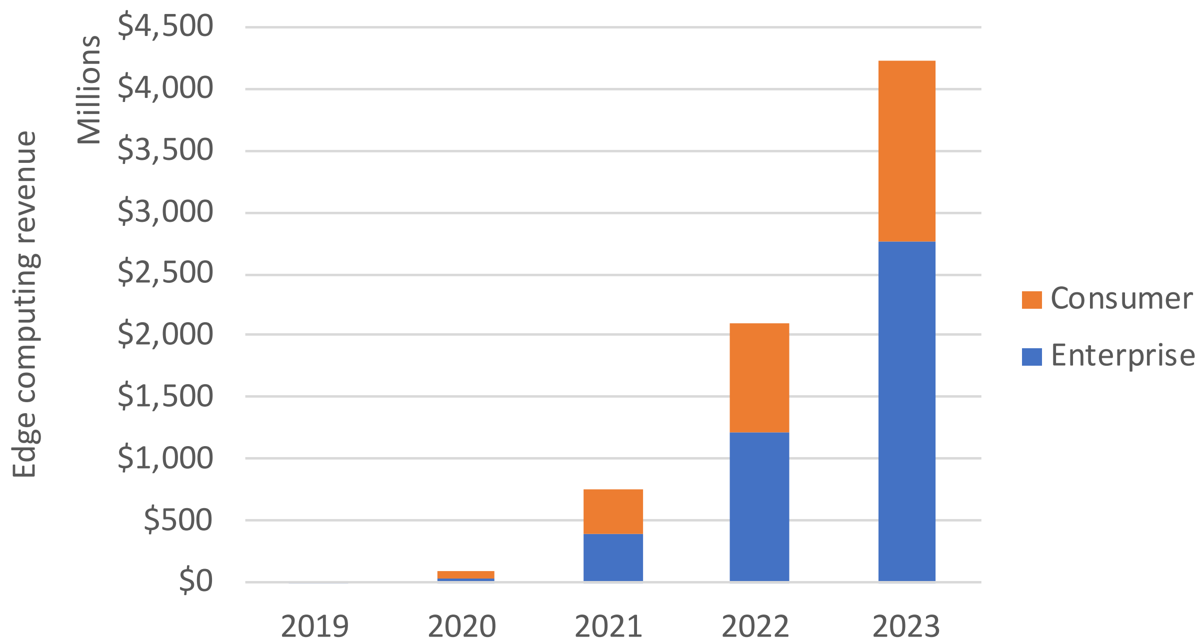



**Figure 12 Edge computing revenue forecast.**

We also forecast that MNOs and enterprises would spend a cumulative of $4.4 billion, through 2023, on edge computing hardware and software.

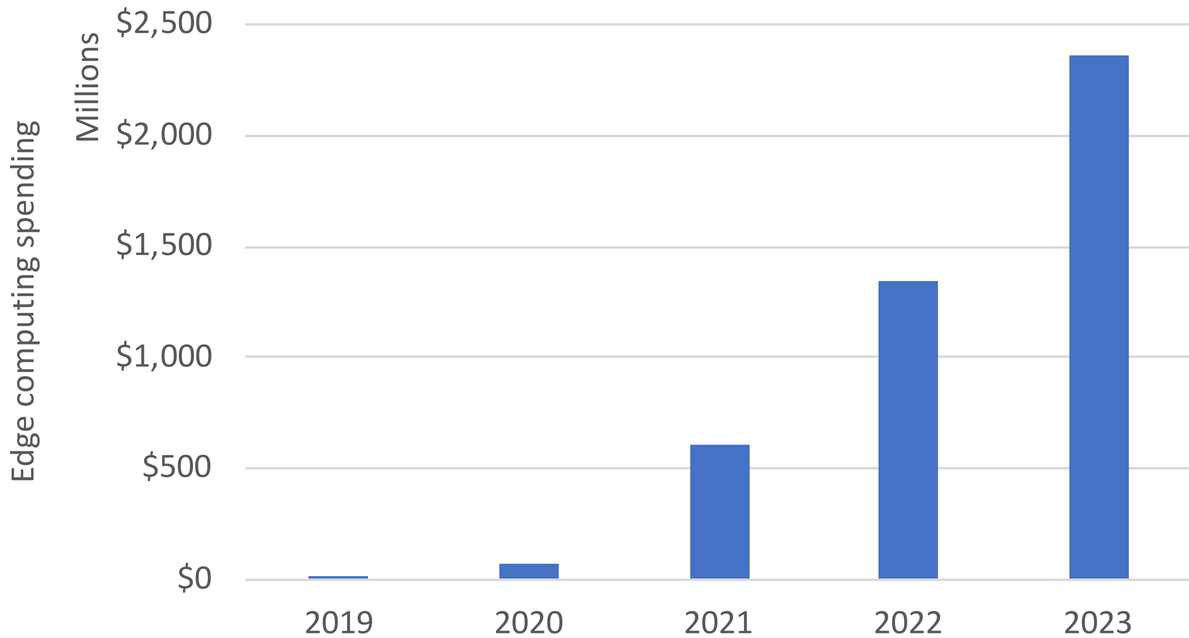



**Figure 13 Edge computing spending forecast.**

The North American market is set to lead the deployment of edge computing, accounting for 59% of cumulative revenue in both enterprise and consumer segments, due to a number of favorable elements:

1- Availability of CBRS 3.5 GHz band which we expect to stimulate private enterprise deployments that leverage edge computing.
2- High penetration of virtualization in networks, and active steps by leading operators to re-architect for higher decentralization.
3- Plans by service providers to deploy 5G technology at a relatively fast pace.
4- High consumer ARPU and a higher willingness to spend on value added services than in other parts of the world.

Japan and Korea would also move at a relatively fast pace, followed by China where we expect IoT applications to positively impact deployment of edge computing. Penetration in Europe would be at slower pace than in other regions primarily due to later deployment of 5G services, although many operators are actively transforming their current 4G networks through virtualization.

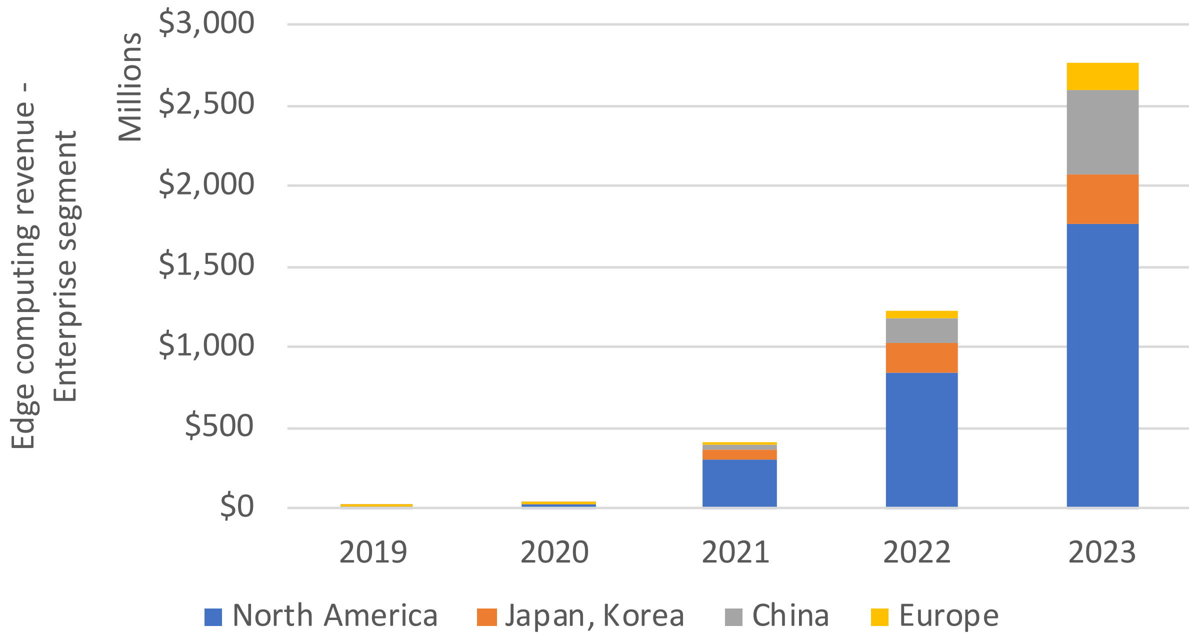



**Figure 14 Forecast of enterprise spending on edge computing by geography.**

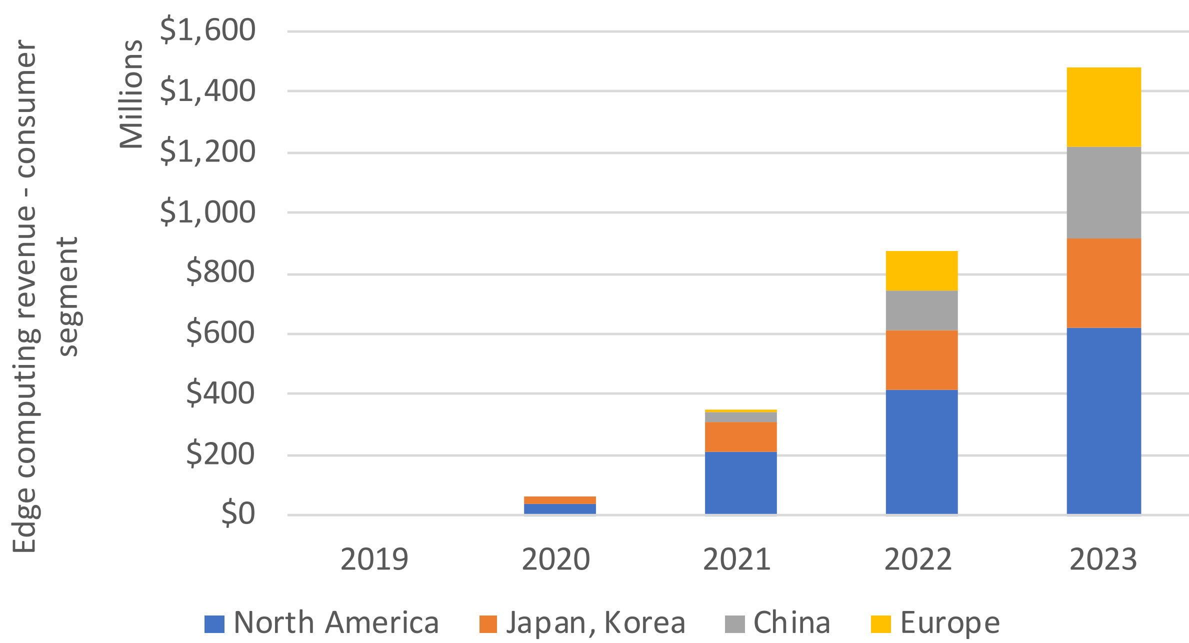



**Figure 15 Forecast of edge computing revenue from consumer segment by geography.**

## Acronyms

| | |
|---|---|
| **4G** | Fourth Generation |
| **5G** | Fifth Generation |
| **AR** | Augmented Reality |
| **ARPU** | Average revenue per user |
| **AWS** | Amazon Web Services |
| **B2B** | Business to business |
| **B2B2C** | Business to business to consumer |
| **B2C** | Business to consumer |
| **CBRS** | Citizen Broadband Radio Service |
| **CORD** | Central Office Re-architected as Data Center |
| **eMBMS** | Enhanced Multimedia Broadcast Multicast Service |
| **ETSI** | European Telecommunications Standards Institute |
| **FCN** | Fog Computing Nodes |
| **HTTPS** | Hyper Text Transfer Protocol with Secure Sockets Layer |
| **IMS** | IP Multimedia Subsystem |
| **IoT** | Internet of Things |
| **ISG** | Industry Specification Group |
| **IT** | Information Technology |
| **LAA** | Licensed Assisted Access |
| **LTE** | Long Term Evolution |
| **M-CORD** | Mobile Central Office Re-architected as Data Center |
| **MEC** | Multi-access Edge Compute |
| **MNO** | Mobile Network Operator |
| **NFV** | Network Function Virtualization |
| **OT** | Operation technology |
| **OTT** | Over the Top |
| **PCRF** | Policy and Charging Rule Function |
| **RAN** | Radio Access Network |
| **RTT** | Round-trip delay time |
| **SCADA** | Supervisory Control and Data Acquisition |
| **TEM** | Telecom Equipment Manufacturers |
| **V2X** | Vehicle to anything |
| **vEPC** | Virtual Enhanced Packet Core |
| **VNF** | Virtual Network Function |
| **VoLTE** | Voice over LTE |
| **VR** | Virtual Reality |
| **WiSUN** | |

## Appendix 1: The OTT Model – AWS as an Example

AWS is one of the best examples of how Cloud players are approaching edge computing with a service called Greengrass for IoT devices. Greengrass allows devices to act locally on the data they generate to minimize the cost of transmitting data to the cloud. Greengrass features a common programming model across devices and cloud so users could develop and test software in the cloud and then deploy it on the edge. Devices can operate offline and quickly process local events, reducing the cost of IoT applications.

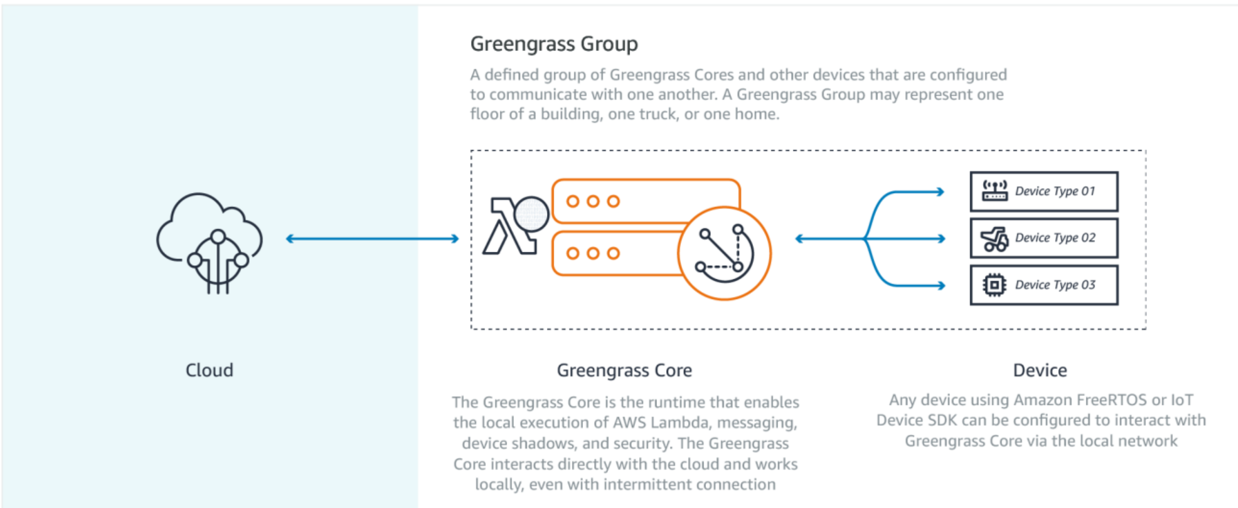



**Figure 16 AWS Greengrass service. [Source: AWS]**

AWS has a total of 114 points of presence including 11 regional edge caches. These are spread into 18 geographic regions worldwide and are subsequently divided into a total of 52 availability zones. Greengrass is available in the following regions, which is set to expand:

- US East (N. Virginia)
- US West (Oregon)
- EU (Frankfurt)
- Asia Pacific (Sydney)
- Asia Pacific (Tokyo)

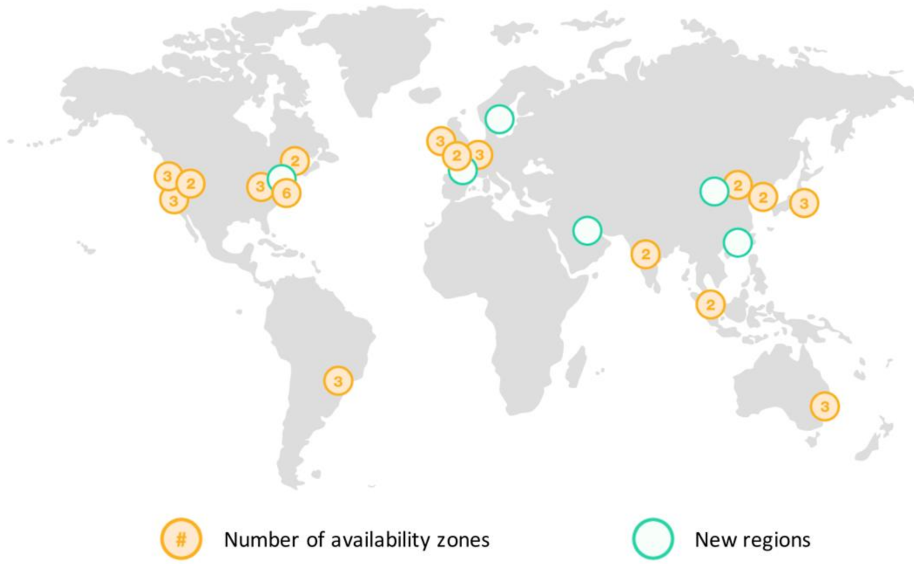



**Figure 17 AWS global infrastructure. [Source: AWS]**

AWS is now a business with $20 billion run rate, having achieved $5.1 billion in revenue in 4Q2017.

# Appendix 2: ETSI MEC Architecture

## Architecture Overview

The ETSI MEC ISG defined a reference architecture (Figure 18) and embarked on specifying interfaces (

Table ***12***) [see Appendix 3: Overview of MEC Interfaces]. Notably, it is not part of the MEC initiative to specify the MEC hosting infrastructure, or the connectivity to the radio access and core network elements (base station or RNC). It is also beyond the MEC initiative to define the interface towards the hosting infrastructure management system. Among the interfaces, the Mp1 interface between the edge application and MEC platform is key for multi-vendor integration. These are important aspects to note in the context of interoperability with other network elements.

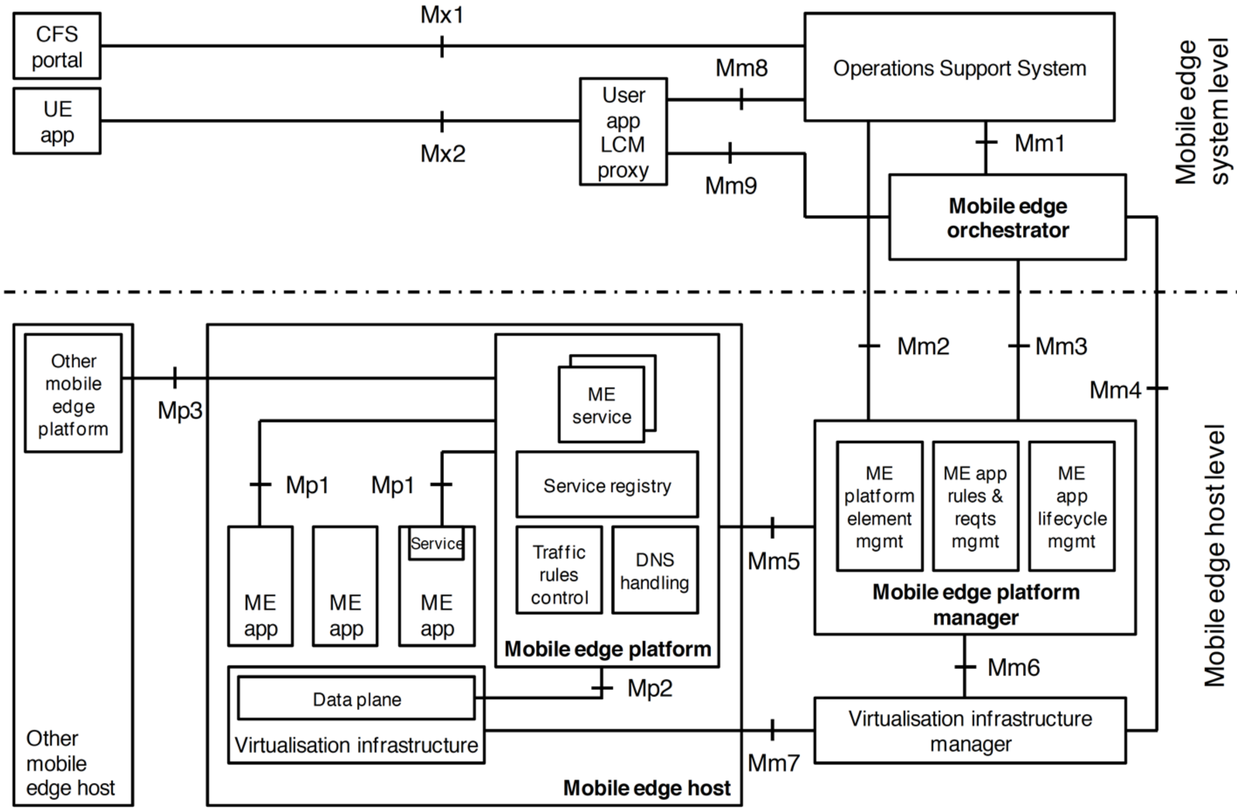



**Figure 18 MEC architecture. [Source: ETSI]**

**Table 12 Interface specification status.**

| Specified Interfaces | Unspecified Interfaces |
|---|---|
| Mm1, Mm2, Mm3, Mm4, Mm6, Mp1, Mp3 | Mm5, Mm7, Mm8, Mm9, Mp2, Mx1 |

The major components of the MEC architecture include:

- Mobile Edge Platform: Sets the policy and configuration rules for forwarding user plane traffic to MEC applications. It also provides a set of services that expose radio network data and other real-time context information to authorized MEC applications.
- Mobile Edge Orchestrator: maintains an overall view of the deployed MEC servers to determine the optimum location for instantiating a MEC application.
- Mobile Edge Platform Manager: responsible for lifecycle management of the MEC applications and management of the MEC Application Platform.
- Virtualized Infrastructure Manager: responsible for managing the resources of virtualized infrastructure, which also includes preparing the infrastructure for running a software image.
- Mobile Edge Applications: MEC-enabled services provided independently of the platform are what make MEC valuable. They must use the MEC application programing interfaces (APIs) and be manageable within the NFV framework.

## MEC Placement in the Network

Unlike 5G, The LTE network does not make provisions for edge computing. Implementing MEC in LTE requires defining an architecture within the constraints of 3GPP standards. Two main network architectures supporting MEC emerge:

1. **MEC on the S1 interface**: The MEC host infrastructure resides between the base station and the core network (S-GW, P-GW and MME) (Figure 19). This architecture fits existing deployments but has critical shortcomings. Mainly, the LTE architecture places functions such as billing, content filtering and legal intercept at the core network. Placing the MEC between the core and the edge requires provisions to restore these critical functions and maintain compliance with the LTE architecture specifications. These provisions are proprietary to the MEC providers and are not governed by ETSI MEC specifications.

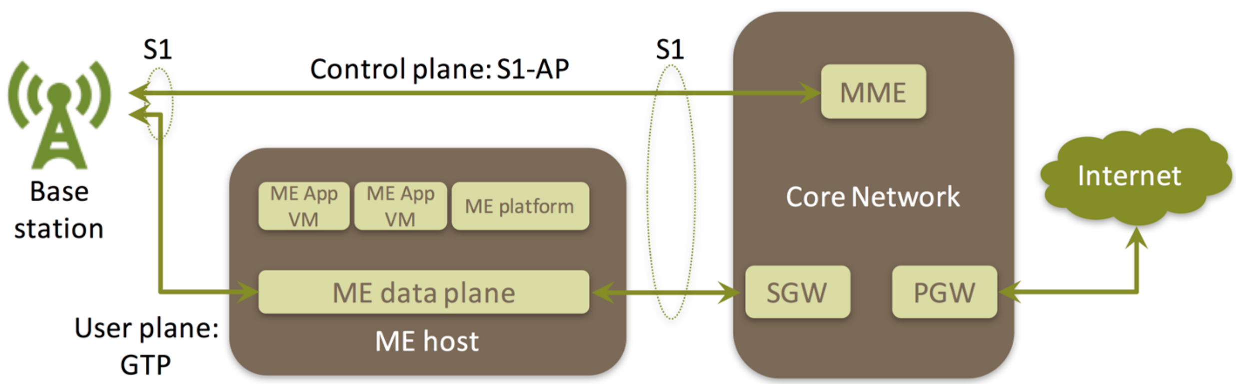



**Figure 19 MEC deployment on the S1 interface.**

2. **MEC on the SGi interface**: It is possible to distribute the core network by pushing elements such as the S-GW and P-GW closer to the edge, although this requires deploying more of these elements than are present in current networks. MEC would

reside on the SGi interface towards the Internet (Figure 20). This architecture accounts for the reservations by many industry players on the S1 MEC architecture.

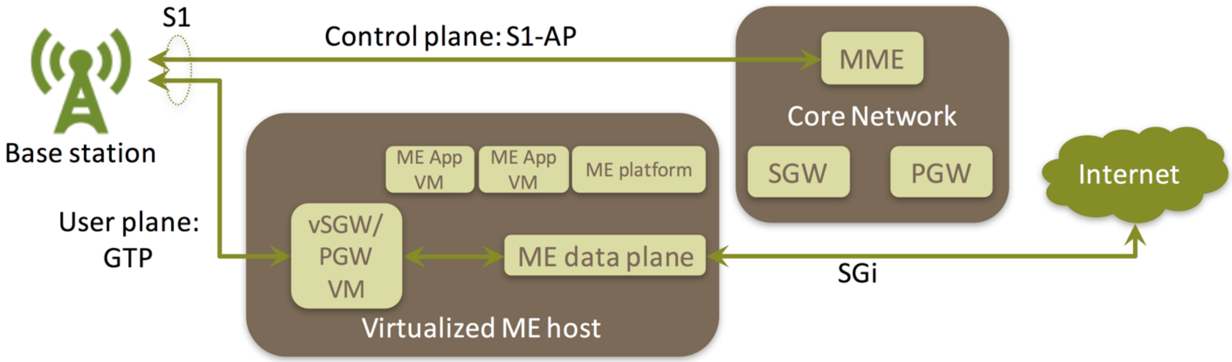



**Figure 20 MEC deployment on the SGi interface.**

A key characteristic is that the MEC architecture is defined to act on user traffic, which could be rerouted or passed through after inspection and/or modification. MEC does not involve control signaling between the core network and base stations. Here, we make the distinction that the control plane over the S1-MME is SCTP-based and the user plane over S1-U is GTP-based. S1 traffic in some cases is also encapsulated in IPSec over the IP backhaul network. The MEC platform should be transparent to GTP traffic which mandates opening a tunnel and re-encapsulating traffic once services are applied via a locally-hosted edge application without impacting the core network. In cases where traffic is encapsulated in IPSec, the MEC platform must additionally support distributed security gateway function.

# Appendix 3: Overview of MEC Interfaces

*Reference Points Related to System-Level Mobile Edge Management*

- Mm1: used for triggering the instantiation and termination of mobile edge applications.
- Mm2: used for configuration, fault and performance management of the mobile edge platform.
- Mm3: used for application lifecycle management, management of application rules & requirements and for keeping track of available mobile edge services
- Mm4: used to manage virtualized resources capacity of the mobile edge host, and to manage application images.
- Mm5: used to perform platform configuration, configuration of application rules and requirements, application lifecycle support procedures, management of application relocation, etc.
- Mm6: used to manage virtualized resources e.g. to realize the application lifecycle management.
- Mm7: used to manage the virtualization infrastructure.
- Mm8: used to handle requests from UE applications to run user applications in the mobile edge system.
- Mm9: used to manage user applications requested by UE applications.

*Reference Points Related to External Entities*

- Mx1: used by the third parties to request permission to run mobile edge applications within the mobile edge system. This reference point is not further specified.
- Mx2 (optional): used by a UE application to request permission to run a user application in the mobile edge system, or to move a user application in or out of the mobile edge system. This reference point is only accessible within the mobile network.

*Reference Points Related to the Mobile Edge Platform*

- Mp1: is the main point of interoperability between platform operator and application provider. It bundles all communication between the mobile edge app and the mobile edge platform.
- Mp2: used to control the data plane in the virtualization infrastructure, including instruction on how to route traffic among applications, networks, and services; and instructions derived from traffic rules provided with the ME application. This reference point is not further specified.
- Mp3: used for control communication between mobile edge platforms.

MEC has chosen the REST API style, with JSON and optional XML data formats to define the APIs.